

Supermarket Model Summary

The purpose of this memo is to summarize the major characteristics of the new Supermarket model, explain how the new model compares with the existing model and provide notes on how stakeholder feedback and existing Portfolio Manager data are incorporated into the final model. *This document is intended for internal reference purposes only.*

New supermarket model

The major characteristics of the new model are as follows:

- Data – The new model is based on the CBECS 1999 and 2003 surveys. Two years worth of data are necessary to generate an adequate sample for analysis.
- Dependent Variable – The dependent unit of analysis is Source Energy Intensity (kBtu/ft²).
- Independent Variables – Table 1 presents the independent variables in the model. These are the operating characteristics that the model adjusts for.
- R² Value – The final model has an R² statistic of 0.514. This value indicates that the model explains 51.4% of the variation in Source EUI among supermarkets. Because this value is computed relative to source EUI, it does not include the explanatory power of square foot. Thus, the value appears artificially low. Re-computing the R² value in units of *source energy* demonstrates that the *model explains 81% of the variation of source energy* in supermarkets. In units of LN(Source Energy), the new model explains 92.8% of variation. This is an excellent result for a statistically based energy model.

The R² is not the only statistic used to evaluate models. Other measures, including individual variable significance levels, the model significance level, and residual plots are also factored into the final decision.

Table 1 Summary of the Final Supermarket Model		
Variable	Model Coefficient	Confidence Level
Intercept	581.112	> 99%
Natural Log of Worker Density	115.601	> 99%
Natural Log of Weekly Operating Hours	125.812	88%
Natural Log of Square Foot	84.972	> 99%
HDD times Percent of the Building Heated	.033	> 99%
CDD times Percent of the Building Cooled	.095	> 99%
Walk-in Refrigeration/Freezer Density	794.388	> 99%
Cooking Density	902.769	83%
Note: - Densities are all expressed on a per 1,000 square foot basis (e.g. number of workers per 1,000 square foot) - Cooking is a yes/no variable. For cooking density, a response of 'no' receives a zero; a response of 'yes' is assigned a value equal to 1/(building size/1000). The density format yields the best result statistically. Physically, it means that the total energy added by cooking is not dependent on the floor area. - Confidence levels of >90% are generally sought in EPA models. However, Operating Hours and Cooking Density are retained based on stakeholder discussions and analysis of PM data.		

Comparison of new model to existing (old) model

The new model is quite different from the existing model. Major differences are as follows:

- Data: the old model was based on CBECS 1992 and 1995 data; the new model uses CBECS 1999 and 2003 data.
 - *More recent data provides a more accurate description of the market*
- New independent variable: the old model was based on LN(Source Energy); the new model is based on Source EUI.
 - *Source EUI is equally robust and easier to explain*
- New independent variables: the new model contains two new variables
 - The percent of the building that is heated
 - The percent of the building that is cooled
 - *The percent heated/cooled enable more accurate and statistically significant adjustments for HDD and CDD*
- Excluded independent variables: the new model does not include three characteristics that are in the existing (old) model:
 - The number of open and closed refrigeration cases
 - This variable was not determined to be statistically significant in any regression formulations (100+ total)
 - Plots of this variable relative to EUI do not exhibit a relationship for either PM or CBECS data (see attached).
 - For both CBECS and PM data, plots of the new ratings relative to the number of these cases do not exhibit any rating bias (see attached).
 - The number of computers and cash registers
 - These variables were not determined to be statistically significant.
 - Their contribution to energy use is likely small relative to characteristics like walk in refrigeration and worker density, both of which are in the model.
 - The number of floors
 - This variable was not determined to be statistically significant
 - The result for this model is consistent with other models (e.g. retail), where the number of floors is not an important factor.
 - *There is no evidence that these operating characteristics are statistically meaningful. This does not mean that they do not use energy. However, the impact of these specific characteristics is accounted for in the other measure of business size within the model (walk in units, worker density, square foot, etc)*
- Significance Levels: Overall, the statistical characteristics of the new model are superior to the existing (old) model.
 - The independent variables are significant with greater levels of confidence.
 - In the existing (old) model, only 4 operating characteristics are significant with 99% confidence, and the other five are *below* 80% confidence.
 - In the new model, 5 out of 7 variables meet the 99% confidence threshold, while the other 2 are still above 80%.
 - In equivalent units of LN(Source Energy), the new model R^2 is greater than the existing (old) model R^2
 - New model $R^2 = 0.928$

- Existing (old) model $R^2 = 0.81$
 - ***The new model is more robust overall.***
- Ratings: The new model provides a more even distribution of ratings across current Portfolio Manager data.
 - With the new model, the average rating in Portfolio Manager is about 56, with approximately 33% of buildings qualifying for the label.
 - Under the existing (old) model, the average rating in Portfolio Manager is about 79, with approximately 74% of buildings qualifying for the label.
 - A graph of these distributions is attached.
 - In a totally random sample, the distribution should be flat. Portfolio Manager is a sample with a known bias; it includes a handful of large organizations and organizations that have superior management practices. As such, it is expected that the distribution will be skewed above 50. However, under the existing model, the number of buildings with high ratings is above what would be expected.
 - ***The new model provides a better (i.e. more uniform) distribution of ratings across the Portfolio Manager Data.***

Incorporation of feedback into the new model

The final model takes into account the statistical analysis of CBECS data, an evaluation of existing data in Portfolio Manager, and also the input from several supermarket stakeholders. EPA held two meetings to gather feedback. The first was with supermarket owner/operators, and was held on March 4, 2008. The second was held on March 5, 2008 with Southern California Edison and PECI, who are working together on a retrocommissioning project of supermarkets. The following items were noted:

- Attendees: Seven supermarket organizations were represented in the first call.
 - Attendance included ENERGY STAR Partners and non-partners
- Percent Heated and Cooled: participants indicated that they could capture these values, noting a preference for being asked for a percent rather than the floor area that is heated or cooled. They also noted that this variable would allow for a more equitable adjustment for HDD and CDD.
 - ***EPA concluded that it is practical and preferable to add these new variables to the model.***
 - ***EPA noted that the variable definition will need to provide guidance on how space that is characterized by walk in refrigeration should be counted.***
- Hours of Operation: participants had a mixed response as to whether they thought hours should be retained in the model. While they acknowledged that many departments close down at night and that much of the equipment (e.g. refrigeration) runs full time, they still identified differences such as lighting. Ultimately there seemed a slight preference for incorporating hours into the model.
 - ***From an analytical standpoint, this variable makes only a marginal difference. It was determined that retaining the variable in the model produces robust and unbiased results, and also meets the expectations of supermarket stakeholders.***

- Cooking: On the whole, supermarket owners and operators felt that the presence of cooking was big differentiating factor among stores. Although most Portfolio Manager buildings have cooking, representatives from PECI also indicated that smaller organizations (perhaps underrepresented in Portfolio Manager) may not have cooking, and therefore that cooking was an important distinction.
 - *Ultimately the models with cooking showed to provide better and more equitable distributions of ratings and these models coincide with the stakeholder expectation. Therefore, this variable was retained in the analysis.*
- Number of Computers and Cash Registers, and the Number of Floors: The participants did not indicate any concern that these variables would be removed from the model.
 - *Computers, Cash Registers, and Number of Floors are not in the model*
- Open and Closed Refrigeration Cases: most stakeholders were surprised to learn that open and closed refrigeration cases did not have a statistically meaningful relationship with respect to energy performance. It was noted that this equipment can account for approximately 45% of the total electrical load.

However, based on the graphs presented by EPA, a few observations became evident to several of the supermarket owners/operators and the representatives from Southern California Edison and PECI. They noted that the graphs were not showing the expected relationship, and that even though the variable is not in a model a supermarket with a greater percentage of closed cases (i.e. more efficient operation) will still rate better under the ENERGY STAR model. Also, the number of open/closed cases may in fact be proportional to other measures of business size, such as the floor area and the number of walk in refrigeration units; these measures are captured in the new model.

- *Based on the statistical evidence, the final model does not include adjustments for open or closed refrigeration cases.*
- *The initial discussions with seven supermarket owner/operators, Southern California Edison, and PECI show that this unexpected result can be explained and justified. Ultimately there are persuasive arguments both from a technical standpoint and also from an operational/programmatic standpoint.*

Attachments

Figure 1: Comparison of EUI and Open and Closed Refrigeration for CBECS and Portfolio Manager Data

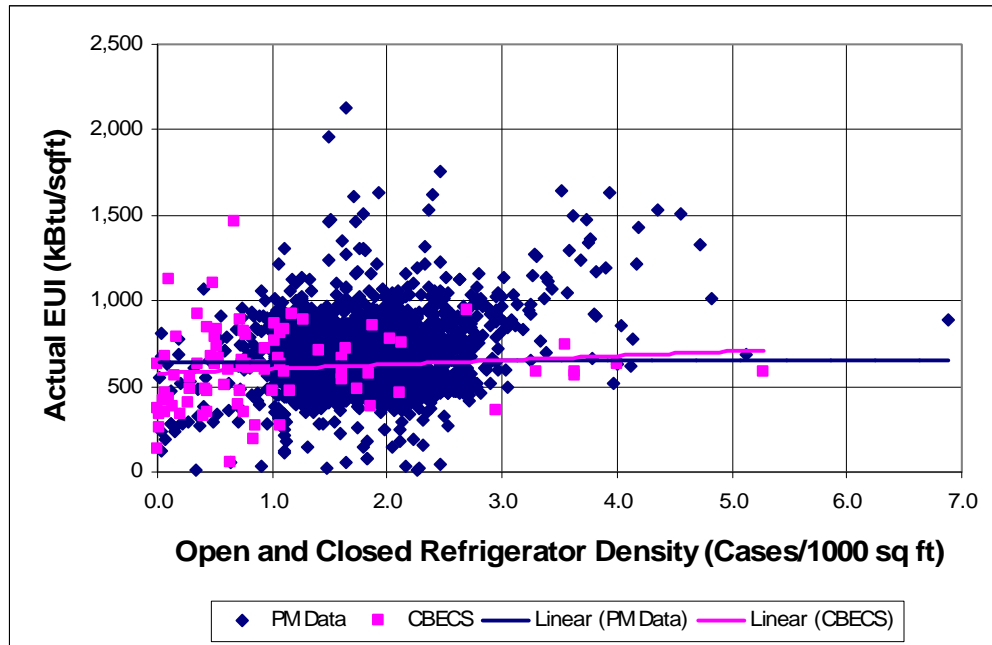


Figure 2: Comparison of Energy Performance Ratings (New Model) with Open and Closed Refrigeration in CBECS and Portfolio Manager Data

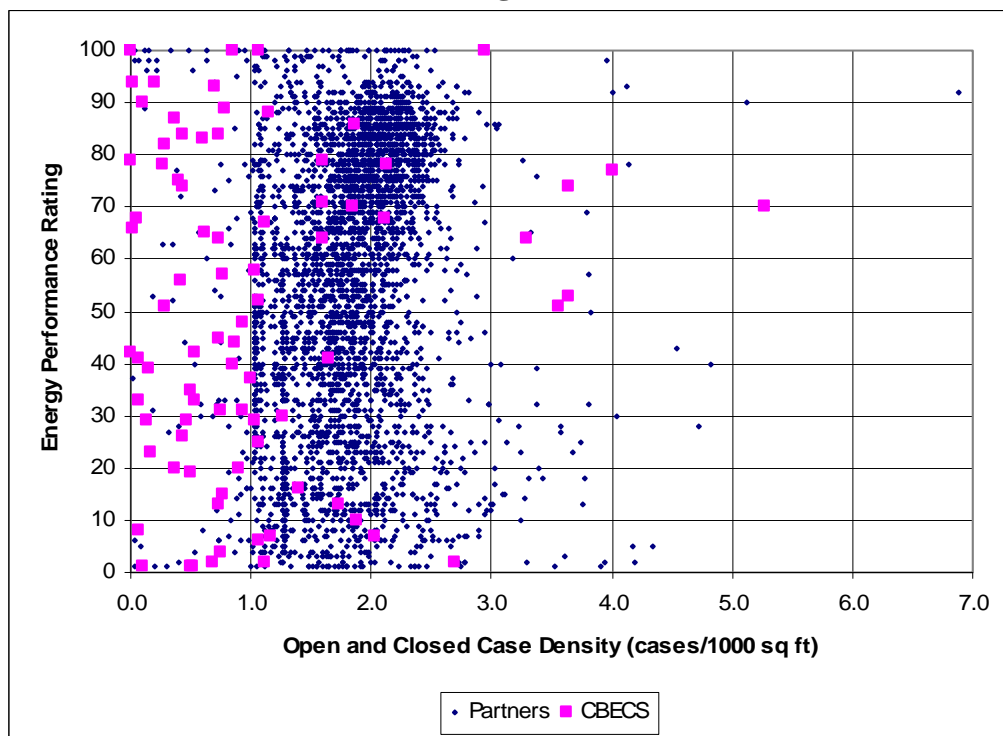


Figure 3: Distribution of Energy Performance Ratings (Portfolio Manager) for the new and the existing (old) models

